## SEQUENCE LISTING

- (1) GENERAL INFORMATION:
  - (i) APPLICANT:
    - (A) NAME: PIERRE FABRE MEDICAMENT
    - (B) STREET: 17, AVENUE JEAN MOULIN
    - (C) CITY: CASTRES
    - (E) COUNTRY: FRANCE
    - (F) POSTAL CODE: 81106
- (ii) TITLE OF INVENTION: IMMUNOGENIC ELEMENT, IMMUNOGENIC 10 AGENT, PHARMACEUTICAL COMPOSITION AND PREPARATION PROCESS.
  - (iii) NUMBER OF SEQUENCES: 75
  - (iv) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk
    - (B) COMPUTER: IBM PC compatible
- 15 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (vi) DATA OF THE LAST APPLICATION:
    - (A) NUMBER OF THE APPLICATION: FR 94 04009
    - (B) FILING DATE: APRIL 06 1994

Information for SEQ ID NO: 1 G2A

20 SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - The Val Lys The Lys Ash The The The The Gln The Gln Pro See Lys Pro The The Lys 5'- ACC GIG AAA ACC AAA AAC ACC ACG ACC ACC CAG ACC CAG CIG AGC AAA CUG ACC AAA 150

Val Pro Cys Ser lle Cys Ser Asn Asn Pro Thr Cys Trp Ala lle Cys Lys Arg lle Pro Asn GTG CTG TGC ; . ATC TGC AGC AAC AAC CTG ACC TGC TGG GGG ATC TGC AAA CGT ATC CTG AAC 192

 Information for SEQ ID NO: 2 G2B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130

Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe ACC CGT ACC AAA AAC CCG CGG AAA AAA CCG AAA GAT GAT TAC CAC TTC GAA GGG TTC AAC TTC 171 173 176 182 186

Val Pro Cys Ser The Cys Gly Asn Asn Gln Leu Cys Lys Ser The Cys Lys Thr The Pro Ser GIG CCC TGC AGC ATC TGC GGC AAC AAC CAG CTG TGC AAA AGC ATC TGC AAAA AGC ATC TGC AAA AGC ATC TGC AAA AGC ATC TGC AAA AGC ATC TGC AAA AGC A

AST LYS PTO LYS LYS LYS PTO THE LIE LYS PTO THE AST LYS PTO THE THE LYS THE THE AST AAC AAA CUE AAA AAG AAA CUE ACC ACC AAA CUE ACC AAA CUE ACC AAA ACC ACC AAC 213

Information for SEQ ID NO: 3 G2A6Cys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - The Val Lys The Lys Ash The The The The Gln The Gln Pro See Lys Pro The The Lys 5'- ACC GIG AAA ACC AAA AAC ACC ACC ACC ACC CAG CCG ACC AAA CCG ACC AAA

Information for SEQ ID NO: 4 G2B6Cys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130 N - The Ala Gin The Lys Gly Arg Die The The See The Gin The Ash Lys Pro See The Lys 5'- ACT COE CHE ACT AMA COE CET ATC ACT ACT ACT ACT CHE ACT AAC AAA COE ACT AMA 150 Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe AGE OFF AGE ANA AND OUR OUR ANA ANA OUR ANA GRIT GRIT THE CHE TITE GRA GIG TITE AND THE 186 182 173 176 171 Val Pro Ser Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser GIG COO AGO AGO AGO GEO AAC AAC CAG CIG TOO AAA AGO ATO AGO AAA ACO ATO CIG AGO 192 Asn Lys Pro Lys Lys Pro Thr Tie Lys Pro Thr Asn Lys Pro Tir Thr Lys Thr Tir Asn ANC ANA COO ANA ANG ANA COO ACC ANC ANA COO ACC ANA COO ACC ANA ACC ACC ANA 230 213 Lys Arg Asp Pro Lys Thr Pro Ala Lys Met Pro Lys Lys Glu Ile Ile Thr Asn - C ANA COST CAST COSS ANA ACCE COSS COSS ANA ASTO COSS ANG ANG CANA ASTO ACCE ANC - 3'

Information for SEQ ID NO: 5 GlACys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187

N - Ser Tile Cys Ser Asn Asn Pro Tir Cys Trp Ala Tile Cys Lys - C

5'- ACC ATC TGC AGC AAC AAC CUG ACC TGC TGG GGG ATC TGC AAA - 3'

Information for SEQ ID NO: 6 G1BCys

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys - C
5'- AGC ATC TGC GGC AAC AAC CAG CTG TGC AAA AGC ATC TGC AAA - 3'

Information for SEQ ID NO: 7 GLA

SEQUENCE TYPE: amino acids and nucleotides

15 SEQUENCE LENGTH: 14 amino acids, 4? nucleotides

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Ser Asn Asn Pro Tim Cys Trip Ala Ile Ser Lys - C
5'- AGC ATC TGC AGC AAC AAC CUG AGC TGC TGG GUG ATC AGC AAA - 3'

Information for SEQ ID NO: 8 G1B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

5 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187
N - Ser Ile Cys Gly Asn Asn Gln Lei Cys Lys Ser Ile Ser Lys - C
5'- AGC ATC TGC GGC AAC AAC CAG CTG TGC AAA AGC ATC AGC AAA - 3'

Information for SEQ ID NO: 9 G1'A

SEQUENCE TYPE: amino acids

10 SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187 N - Ser Ile Asp Ser Asn Asn Pro Thr Crn Trp Ala Ile Cys Lys - C Information for SEQ ID NO: 10 Gl'B

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187 N - Ser Ile Asp Gly Asn Asn Gln Leu Orn Lys Ser Ile Cys Lys - C

Information for SEQ ID NO: 11 G1'A $\delta$ C

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 14 amino acids

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187 N - Ser Ile Asp Ser Asn Asn Pro Thr Orn Trp Ala Ile Ser Lys - C

Information for SEQ ID NO: 12 G1'BδC

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187 N - Ser Ile Asp Gly Asn Asn Gln Leu Om Lys Ser Ile Ser Lys - C Information for SEQ ID NO: 13 P40

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 335 amino acids, 1005 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

P40 = N -Ala Pro Lys Asp Asn The Trp Tyr Ala Gly Gly Lys Leu Gly Trp Ser GCT GGG AAA GAT AAC AGC TGG TAT GCA GGT GGT AAA CIG GGT TGG TGC

GIN TYP HIS ASP THE GIY PARE GIY ASP GIY PARE GIN ASP ASP ASP GIY PARE GIN THE CASE ARE ASE GIF COST THE CASE ARE ARE ASP COST COST OFF

Tym Leu Cily Phe Cilu Met Cily Tym Asp Tup Leu Cily Arg Met Ala Tym Lys Cily The circ Cir Tirc Gra arg Cir Tai Gre Tics circ Cir Cir Arg Cea Tai ara Circ 71

Ser Val Asp Asn Gly Ala Phe Lys Ala Gln Gly Val Gln Leu Thr Ala Lys Leu ACC GTT GAC AAC GGT GCT TTC AAA GCT CAG GCC GTT CAG CTG ACC GCT AAA CTG 89

GIY TYP PRO ILE THE ASP ASP LEU ASP ILE TYP THE ANG LEU GLY GLY MEE VAL GET TAC COE ATC ACT GAC GAT CIG GAC ATC TAC ACC GET CIG GEC GEC ATG GIT 107

The Arg Ala Asp Ser Lys Gly Ash Tyr Ala Ser Thr Gly Val Ser Arg Ser Glu TGG GGC GGC GGC TGC AAA GGC AAC TAC GGT TGT AGC GGC GTT TGC GGT AGC GAA 125

His Asp The Gly Val See Pro Val Pre Ala Gly Gly Val Glu Trp Ala Val The OAC GAC ACT GGT GGT GGA GAG TGG GGT ACT

Arg Asp Ile Ala Thr Arg Leu Glu Tyr Gln Trp Val Asn Asn Ile Gly Asp Ala GGT GAC ATC GGT ACC GGT CTG GAA TAC GAG TGG GTT AAC AAC ATC GGC GAC GGG

GLY THE Val. GLY THE ARG PRO ASP ASS GLY MET LEU SEE LEU GLY Val. SEE TYE GET ACT GIG GET ACT GET GET GET GET ACT GET

197 Pro Glu Val Ala Thr Lys His Phe Thr Leu Lys Ser Asp Val Leu Phe Asn Phe CUE GAA GUE CET ACE AAG CAC TUC ACE CUE AAG TET GAE GUT CUE TUC AAC TUC 215 Asn Lys Ala Thr Leu Lys Pro Glu Gly Gln Gln Ala Leu Asp Gln Leu Tyr Thr AAC AAA GOT ACC CTG AAA COO GAA GOT CAG CAG COT CTG GAT CAG CTG TAC ACT 233 Gin Leu Ser Asn Met Asp Pro Lys Asp Gly Ser Ala Val Val Leu Gly Tyr Thr CAG CITG ACC AAC ATTG GAT COT AAA GAC GET TOO COT GIT CITG GEC TAC ACC 251 Asp Arg Ile Gly Ser Glu Ala Tyr Asn Gln Gln Leu Ser Glu Lys Arg Ala Gln GAC COC ATC COST TOC GAA COT TAC AAC CAG CAG CAG TOT GAG AAA COST CAG 269 Ser Val Val Asp Tyr Leu Val Ala Lys Gly Ile Pro Ala Gly Lys Ile Ser Ala TOO GIT GIT GAC TAC CIG GIT GOT AAA GOO AIC CIG GOT GOO AAA AIC TOO GOT 287 Arg Gly Met Gly Glu Ser Asn Pro Val Tir Gly Asn Tir Cys Asp Asn Val Lys CCC CCC ATG CGT GAA TCC AAC CCG GTT ACT CCC AAC ACC TGT GAC AAC GTG AAA 305 Ala Arg Ala Ala Leu Ile Asp Cys Leu Ala Pro Asp Arg Arg Val Glu Ile Glu 335 323 Val Lys Gly Tyr Lys Glu Val Val Tir Gln Pro Ala Gly TER - C GIT AAA GGC TAC AAA GAA GIT GIA ACT CAG GGG GGT TAA - 3'

Information for SEQ ID NO: 14 G2A $\delta$ CF

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130 N - The Val Lys The Lys Asn The The The Cln The Cln Pro See Lys Pro The The Lys 5'- ACE GIG AAA ACE AAA AAC ACE ACE ACE ACE CAG CAG CAG CAG ACE AAA COG ACE ACE ACE AAA 170 168 165 163 150 Gin Arg In Asn Lys Pro Pro Asn Lys Pro Asn Asn Asp Ser His Ser Glu Val Ser Asn Ser OPG CET CAG AAC AAA CCC CCC AAC AAA CCC AAC GAT TOO CAT TOO GAA GTG TOO AAC TOO 186 182 176 171 Val Pro Ser Ser Ile Cys ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn GIG CUI MEC MEC MIC TIEC MEC MAC MAC CUI MEC TIEC TIEC GUI MIC MEC MAA CUIT MIC CUIG MAC 192 Lys Lys Pro Gly Lys Lys The The The Lys Pro The Lys Lys Pro The He Lys The The Lys AAA AAA COG GGC AAA AAA ACC ACG ACC AAA COG ACC AAA AAA COG ACC TIC AAA ACC ACC AAA 213 230

Lys Asp His Lys Pro Gln The The Lys Pro Lys Glu Val Pro The The Lys Pro - C
AAA GAT CAT AAA CUG CAG ACC ACC AAA CUG AAA GAA GUG ACC ACC AAA CUG - 3'

Information for SEQ ID NO: 15 G4A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Cys Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys - C
5'- GIG CUS TGC AGC AGC AGC AGC AGC AGC AGC TGG GUS AGC TGC GGG GUS AGC TGC AAA - 3'

Information for SEQ ID NO: 16 G4AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187 N - Val Pro Ser Ser Ile Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys - C

5' - GIG CCG AGC AGC AGC AGC AGC AAC CCG ACC TGC TGG GCG ATC AGC AAA - 3'

Information for SEQ ID NO: 17 G4B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Cys Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys - C5' -GIG CCC TGC AGC ATC TGC GGC AAC AAC CAG CTG TGC AAA AGC ATC TGC AAA - 3'

Information for SEQ ID NO: 18 G4B&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 42 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 182 186 187
N - Val Pro Ser Ser Ile Cys Gly Asn Asn Gln Leu Cys Lys Ser Ile Ser Lys - C
5' -GIG CCC AGC AGC AGC AGC CGC CGC CGC TGC AAA AGC ATC AGC AAA - 3'

Information for SEQ ID NO: 19 G4'A

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187 N - Val Pro Asp Ser Ile Asp Ser Asa Asa Pro Thr Ozn Trp Ala Ile Ozn Lys - C WO 95/27787

Information for SEQ ID NO: 20 G4'A&C

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

186 187 182 176 173 171 N - Val Pro Ser Ser Ile Asp Ser Asn Asn Pro Thr Om Trp Ala Ile Ser Lys - C

Information for SEQ ID NO: 21 G4'B

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 17 amino acids

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

186 187 182 176 173 171 N - Val Pro Asp Ser Ile Asp Gly Asn Asn Gln Leu Orn Lys Ser Ile Orn Lys - C

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

182 176 173 171 N - Val Pro Ser Ser Ile Asp Gly Asn Asn Gln Leu Orn Lys Ser Ile Ser Lys - C Information for SEQ ID NO: 23 G200A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin The Gin Pro Ser Lys Pro The The Lys Gin Arg Gin Asn Lys Pro Pro Asn 5'- CG ACC CAG CCG ACC AAA CCG ACC AAA CAG CGT CAG AAC AAA CCG ACC AAC IS8

173
176
1ys Pro Asn Asn Asp Pre His Pre Glu Val Pre Asn Pre Val Pro Cys See Lie Cys

ANA CUE AND AND AND THE HIS HE GIT VAI HE AND HE VAI HE VAI HE CHYS

ANA CUE AND AND GRE THE CREE THE GREATH AND THE GITS CUE HER AND THE

1872

1886

Ser Ash Ash Pro The Cys Trp Ala Ile Cys Lys Arg Ile Pro Ash Lys Lys Pro Gly
ACC AAC AAC ACC TEC TEC ACC ATC TEC AAA CET ATC ACC AAA AAA AAA CEC ACC
196 200

Lys Lys The The The - C AAA AAA ACC ACG ACC - 3'

Information for SEQ ID NO: 24 G198A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - GIN The GIN Pro See Lys Pro The The Lys GIN Arg GIN Asn Lys Pro Pro Asn 5'- CAG ACC CAG CTG AGC AAA CTG ACC ACC AAA CAG CGI CAG AAC AAA CTG CTG AAC 158

Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys AAA CUS AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC GIG CUS TGC AGC ATC TIC 127 182 186

SET ASI ASI Pro The Cys Trp Ala Ile Cys Lys Arg Ile Pro Asi Lys Lys Pro Gly ACC AAC AAC COE ACC TOE GOE ATC TOE AAA COE ATC COE AAC AAA AAA COE GOE 196 198

Lys Lys Thr - C

AAA AAA ACE - 3'

Information for SEQ ID NO: 25 G196A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin Thr Gin Pro Ser Lys Pro Thr Thr Lys Gin Arg Gin Asn Lys Pro Pro Asn .

5'- CAG ACC CAG CCG AGC AAA CCG AGC AAA CAG CGT CAG AAC AAA CCG CCG AAC 173 176

Lys Pro Asr Asr Asp Pre His Pre Glu Val Pre Asr Pre Val Pro Cys Ser Ile Cys
AAA CCG AAC AAC GAT TIC GAT TIC GAA GIG TIC AAC TIC GIG CCG TGC AGC ATC TGC
177 186

Ser Asn Asn Pro The Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro Gly ACC AAC AAC CCG ACC TGC TGC GCG ATC TGC AAA CGT ATC CCG AAC AAA AAA CCG GCC 196

Lys - C

AAA - 3'

Information for SEQ ID NO: 26 G194A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - GIN The GIN Pro See Lys Pro The The Lys GIN Arg GIN Asn Lys Pro Pro Asn 5'- CAG ACE CAG GUE AGE AAA CUE ACE ACE AAA CAG GUE CAG AAC AAA CUE GUE AAC 158

Lys Pro Asn Asn Asp Pre His Pre Glu Val Pre Asn Pre Val Pro Cys See Lie Cys AAA CUE AAC GAE TIC CAE TIC CAA GUE TIC AAC TIC GUE GUE TIC AGE AUC TIC 177

182

186

194

Ser Asn Asn Pro The Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro - C
ACC AAC AAC COE ACC TOE TOE GOE ATC TOE AAA COE ATC COE AAC AAA AAA COE - 3'

Information for SEQ ID NO: 27 G192A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin The Gin Pro See Lys Pro The The Lys Gin Arg Gin Asn Lys Pro Pro Asn . 158 173 176 Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys AAA CUE AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC GIG CUE IGC AGC AIC IGC 177 182 186 Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys - C AGC AAC AAC COG ACC TGG TGG GGG ATC TGC AAA CGT ATC COG AAC AAA - 3'

Information for SEQ ID NO: 28 G6A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single 10

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin The Gin Pro See Lys Pro The The Lys Gin Arg Gin Asn Lys Pro Pro Asn 5'- CPG ACC CPG CCG ACC AAA CCG ACC ACC AAA CPG CGT CPG AAC AAA CCG CCG AAC 173 176 158 Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys AAA CCG AAC AAC GAT TIC CAT TIC GF : GIG TIC AAC TIC GIG CCG TGC AGC ATC TGC 182 186 190 -177 Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro - C ACC AAC AAC COG ACC TOC TOG COG ATC TOC AAA COT ATC COG - 3'

Information for SEQ ID NO: 29 G7A

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

158

MOLECULE TYPE: protein

173

N - Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile 5' - AAA CUE AAC AAC GAT TIC CAT TIC GAA GUG TIC AAC TIC GUG CUE TOC AGC AUC

176 182 186 190

Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro - C

TOO AGO AAC AAC COG ACO TOO TOO TOO ATC TOO AAA COST ATC COG - 3'

Information for SEQ ID NO: 30 G200A&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Cin The Cin Pro See Lys Pro The The Lys Cin Arg Cin Asn Lys Pro Pro Asn

5'- CAG ACC CAG CCG AGC AAA CCG ACC AAC CAG CGT CAG AAC AAA CCG CCG AAC

158 173 176

Lys Pro Asn Asn Asp Pre His Pre Glu Val Pre Asn Pre Val Pro Ser Ser Ile Cys

AAA COG AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC GIG COG AGC AGC AGC TIC TGC

177 182 186

Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys Lys Pro Gly

ACC AAC AAC COE ACC TEE TEE GOE ATC AGC AAA CET ATC COE AAC AAA AAA COE GOC

196 200

Lys Lys Tim Tim Tim - C

AAA AAA ACC ACG ACC - 3'

Information for SEQ ID NO: 31 G198A&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys AAA CUE AAC AAC GAT TIC GAT TIC GAA GIG TIC AAC TIC GIG CUE AGC AGC ATC TIC 177 182 186

Ser Ash Ash Pro The Cys Top Ala Ile Ser Lys Arg Ile Pro Ash Lys Lys Pro Gly ASC AAC AAC CUG ACC TGC TGG GGG ATC AGC AAA GGT ATC CUG AAC AAA AAA CUG GGC 196 198

Lys Lys Tir - C

AAA AAA ACE - 3'

Information for SEQ ID NO: 32 G196AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin The Gin Pro See Lys Pro The The Lys Gin Arg Gin Asn Lys Pro Pro Asn 5'- CAG ACC CAG CCG ACC AAA CCG ACC AAA CAG CCG CAG AAA CCG ACC AAA

Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys
AAA CUG AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC GIG CUG AGC AGC ATC TIC
177 182 · 186

Ser Asn Asn Pro The Cys Top Ala Ile Ser Lys Arg Ile Pro Asn Lys Lys Pro Gly
AGC AAC AAC CUG ACC TIC TIC GUG ATC AGC AAA CUT ATC CUG AAC AAA AAA CUG GGC
196

Lys - C AAA - 3'

Information for SEQ ID NO: 33 G194A&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

Information for SEQ ID NO: 34 G192AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin The Gin Pro Ser Lys Pro The The Lys Gin Arg Gin Ash Lys Pro Pro Ash 5'- CAG ACE CAG GUE AGE AAA GUE ACE ACE AAA CAG GUT CAG AAC AAA GUE GUE AAC

176 173 158 Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys AAA COG AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC GIG COG AGC AGC AGC ATC TGC 192 182 186 177 Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro Asn Lys - C AGC AAC AAC COO ACC TOO TOO TOO GOO ATC AGC AAA COT ATC COO AAC AAA - 3'

Information for SEQ ID NO: 35 G6AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 50 amino acids, 150 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Gin The Gin Pro See Lys Pro The The Lys Gin Arg Gin Asn Lys Pro Pro Asn 5'- ONG ACT CAG CCG ACC AAA CCG ACC ACC AAA CAG CGT CAG AAC AAA CCG CCG AAC 176 158 Lys Pro Asn Asn Asp Rhe His Rhe Glu Val Rhe Asn Rhe Val Pro Ser Ser Ile Cys AAA CCG AAC AAC GAT TIC CAT TIC GAA GIG TIC AAC TIC GIG CCG AGC AGC AIC TGC 182 190 186 177 Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro - C AGE AAC AAC COG ACC TOC TOG GCG ATC AGC AAA CGT ATC CCG - 3'

Information for SEQ ID NO: 36 G7AδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

173 158 N - Lys Pro Asn Asn As, The His The Glu Val The Asn The Val Pro Ser Ser Ile 5'- AND COG AND AND GRO THE CAT THE GRA GIG THE AND THE GIG COG AGE AGE AND

182 190 176 186 Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Ser Lys Arg Ile Pro - C TEC AGE AND AND COE ACE TEE TEE COE ATE AGE AND CET ATE COE - 3'

Information for SEQ ID NO: 37 G200B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Tir Gin Tir Asn Lys Pro Ser Tir Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 5'- AGC AGC CAG ACC AAC AAA COG AGC AGC AAA AGC CGT AGC AAA AAC COG CGG AAA AAA COG 160 173 176

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CCC IGC AGC AIC IGC GGC AAC AAC CAG 182 186 200

Leu Cys Lys Ser Ile Cys Lys Tir Ile Pro Ser Asn Lys Pro Lys Lys Pro Tir Ile- C CIG TOC AAA AGC ATC TOC AAA ACC ATC COG AGC AAC AAA COG AAA AAG AAA COG ACC ATC- 3'

Information for SEQ ID NO: 38 G198B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gin Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 5'- AGC ACC CAG ACC AAC AAA COG AGC ACC AAA AGC CGT AGC AAA AAC COG COG AAA AAA COG

173 176 160

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val i o Cys Ser Ile Cys Gly Asn Asn Gln AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG GCC IGC AGC AIC IGC GGC AAC AAC CAG 182 186 198

Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro - C CIG TOE AAA AGE ATC TOE AAA ACE ATC CCG AGE AAC AAA CCG AAA AAG AAA CCG - 3' Information for SEQ ID NO: 39 G196B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 5'- AGC ACC CAG ACC AAA CCG ACC AAA AGC CGT AGC AAA AAC CCG AAA AAA CCG 160

Lys Asp Asp Tyr His Five Glu Val Five Asn Five Val Pro Cys Ser Ile Cys Gly Asn Asn Gln
AAA GAT GAT TAC CAC TIC GAA GTG TTC AAC TTC GTG CCC TGC AGC ATC TGC GGC AAC AAC CAG
182 186

Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Ash Lys Pro Lys Lys - C CIG TGC AAA AGC ATC TGC AAA AGC ATC CGG AGC AAC AAA CUG AAA AAG - 3'

Information for SEQ ID NO: 40 G194B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gin Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 5'- ACC ACC CAG ACC AAA COG ACC ACC AAA ACC CGT ACC AAA AAC COG COG AAA AAA COG

160 173 176

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn Asn Gln AAA GAT GAT TAC CAC TTC GAA GTG TTC AAC TTC GTG CCC TGC AGC ATC TGC GGC AAC AAC CAG 182 186 194

Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro - C CIG TGC AAA AGC ATC TGC AAA AGC ATC GGG AGC AAC AAA GGG · 3' Information for SEQ ID NO: 41 G192B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 53 amino acids, 159 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 5'- ACC ACC CAG ACC AAA COG ACC AAA ACC CGT ACC AAA AAC CCG CCG AAA AAA CCG 160

Lys Asp Asp Tyr His Rhe Glu Val Rhe Ash Rhe Val Pho Cys Ser Ile Cys Gly Ash Ash Gln AAA GAT GAT TAC CAC TTC GAA GTG TTC AAC TTC GTG CCC TGC AGC ATC TGC GGC AAC AAC CAG 1822 186 192

Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn - C

Information for SEQ ID NO: 42 G6B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

Lys Asp Asp Tyr His Rhe Glu Val Rhe Ash Phe Val Pro Cys Ser Ile Cys Gly Ash Ash Gln AAA GAT GAT TAC CAC TTC GAA GTG TTC AAC TTC GTG CCC TGC AGC ATC TGC GGC AAC AAC CAG 182 190

Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro - C CIG TGC AAA AGC ATC TGC AAA AGC ATC CIG - 3° Information for SEQ ID NO: 43 G7B

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

158

N - Lys Pro Lys Asp Asp Tyr His Pre Cilu Val Pre Asr Pre Val Pro Cys Ser Ile Cys Cily

5'- AAA CII AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CIC TIC ACC AIC TIC GIG

182

186

190

Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro - C AAC AAC CAG CIG TGC AAA AGC ATC TGC AAA AGC AIC CIG - 3'

Information for SEQ ID NO: 44 G200B $\delta$ C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser The Gin The Ash Lys Pro Ser The Lys Ser Arg Ser Lys Ash Pro Pro Lys Lys Pro 5'- ACC ACC AAC AAC AAC ACC ACC ACC AAC AAC ACC ACC ACC ACC ACC ACC ACC AAC AAC ACC AC

Lys Asp Asp Tyr His Pre Glu Val Pre Asn Pre Val Pro Ser Ser Lie Cys Gly Asn Asn Ghi AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CIC ACC ACC ACC ACC CAC AAC AAC CAG 182 186

Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro Thr Ile- C CTG TCC AAA ACC ATC ACC AAA ACC ATC CCG ACC AAA CCG AAA AAG AAA CCG AAA CCG ACC ATC- 3' Information for SEQ ID NO: 45 G198BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gin Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 5'- ACC CCG ACC AAC AAA CUE ACC ACC AAA AAC CUE AAA AAC CUE CUE AAA AAA CUE 160 Lys Asn Asn Thr His Pro Gin Val Pro Asn Pro Val Pro Ser Ser Lie Cys Gily Asn Asn Gin

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CCC AGC AGC AGC AGC AAC AAC CAG 182 186

Leu Cys Lys Ser Ile Ser Lys Thr Ile Pro Ser Asn Lys Pro Lys Lys Pro - C CIG TGC AAA AGC ATC AGC AAA AGC ATC GGG AGC AAC AAA GGG AAA AAG AAA GGG - 3'

Information for SEQ ID NO: 46 G198B&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gin Thr Aen Lys Pro Ser Thr Lys Ser Arg Ser Lys Aen Pro Pro Lys Lys Pro 5'- AGC ACC CAG ACC AAA CUG ACC AAA ACC CGT AGC AAA AAC CUG CGG AAA AAA CUG 173 176

160

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln

AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CCC AGC AGC AGC AGC CAC

182

186

Less lys Ser De Ser Lys Thr De Pro Ser Ash Lys Pro Lys Lys - C CIG TCC AAA ACC ATC ACC AAA ACC ATC CCG AGC AAC AAA CCG AAA AAG - 3' Information for SEQ ID NO: 47 G194BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

LEU CYS LYS SET LIE SET LYS THE LIE PRO SET ASI LYS PRO - C CIG TGC AAA AGC ATC AGC AAA AGC ATC CGG AGC AAC AAA CGG - 3'

Information for SEQ ID NO: 48 G192B&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 53 amino acids, 159 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gln Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Liys Lys Pro 5'- AGC ACC CAG ACC AAA CCG AGC AGC AAA AGC CGT AGC AAA AAC CCG CGG AAA AAA CCG 160

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CCC AGC AGC AGC AGC AAC AAC CAG 1822 186 192

LEU Cys Lys Ser Ille Ser Lys Thr Ille Pro Ser Asn - C CIG TCC ANA ACC ATC ACC ANA ACC ATC CCG ACC AAC - 3' Information for SEQ ID NO: 49 G6B&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Ser Thr Gin Thr Asn Lys Pro Ser Thr Lys Ser Arg Ser Lys Asn Pro Pro Lys Lys Pro 51 - AGC ACT CAG ACT AAC AAA COT AGC ACT AAA ACT CET AGC AAA AAC COT COT AAA AAA COT 173 176 160

- 58 -

Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly Asn Asn Gln AAA GAT GAT TAC CAC TIC GAA GIG TIC AAC TIC GIG CCC AGC AGC ATC TGC GGC AAC AAC CAG 186 182

Les Cys Lys Ser Ile Ser Lys Thr Ile Pro - C CIG TGE AAA AGE ATC AGE AAA ACE ATC CCG - 3'

Information for SEQ ID NO: 50 G7BδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single 10

TOPOLOGY: linear

MOLECULE TYPE: protein

173 158

N - Lys Pro Lys Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Ser Ser Ile Cys Gly 5'- AND COS AND GAT GAT THE CHE THE GAD GIG THE AND THE GIG COE AGE AGE ATE THE GAD.

> 182 186

Asn Asn Gin Lau Cys Lys Ser Ile Ser Lys Tir Ile Pro - C

AAC AAC CAG CTG TGC AAA AGC ATC AGC AAA ACC ATC CCG - 3'

Information for SEQ ID NO: 51 G2V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Gin Ash Arg Lys Ile Lys Gly Gin Ser Thr Leu Pro Ala Thr Arg Lys Pro Pro Ile Ash 5'- CAA AAC AGA AAA ATC AAA GGT CAA TCA ACA CTA CCA GCC ACA AGA AAA CCA CCA ATT AAT 150

Pro Ser Gly Ser Ile Pro Pro Glu Ash His Gln Asp His Ash Ash Phe Gln Thr Leu Pro Tyr CCA TCA GGA ACC ATC CCA GGA AAC CAT CAA GGC CAC AAC AAC TIC CAA ACA CIC CCC TAT 171 173 176 182 186

Val Pro Cys Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu GIT CCC TGC AGT ACA TGT GAA GGT AAT CIT GCA TGC TTA TCA CIC TGC CAT ATT GAG AGG GAA 192

Arg Ala Pro Ser Arg Ala Pro Thr Ile Thr Leu Lys Lys Thr Pro Lys Pro Lys Thr Thr Lys AGA CCA ACC AGA CCA ACC ACC ACC ACC CCC AAA AAG ACA CCA AAA CCA AAA ACC ACA AAA 213

Lys Pro The Lys The The His His Arg The See Pro Glu The Lys Leu Gln - C AG CCA ACC AGA ACA ACC CAT CAC AGA ACC AGA ACC AGA ACC AGA ACC AGA CTG CAA - 3'

Information for SEQ ID NO: 52 G2VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 101 amino acids, 303 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

130

N - Gin Ash Arg Lys Cie Lys Cily Gin Ser Thr Leu Pro Ala Thr Arg Lys Pro Pro Ile Ash 5'- CAA AAC AGA AAA AIC AAA GGT CAA TCA ACA CIA CCA GCC ACA AGA AAA CCA CCA ATT AAT 150

Pro Ser Cly Ser The Pro Pro Clu Asn His Cln Asp His Asn Asn Fre Cln Thr Leu Pro Tyr COR TOR GOR ACC ATC COR COR GRA ARC COST CRA GRC CRC ARC ARC TIC CRA ACA CITC CCC TRIT 186

182 176 173 171

Val Pro Ser Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu GIT COE AGE AGI AGA TGI GAA GGI AAT CIT GGA TGE TIA TGA CITC AGE GAT AIT GAG AGE GAA 192

Arg Ala Pro Ser Arg Ala Pro Tir Ile Tir Leu Lys Lys Tir Pro Lys Pro Lys Tir Tir Lys AGA CCA CCA AGC AGA CCA CCA ACA ATC ACC CTC AAA AAG ACA CCA AAA CCA AAA ACC ACA AAA 230 213

Lys Pro The Lys The The His His Arg The See Pro Glu The Lys Leu Gln - C ANG COR ACE ANG ACA ACA ATC CAT CAC ACA ACE ACE COR GAR ACE ANA CTG CAR - 3'

Information for SEQ ID NO: 53 G200V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His 5'- CIA CIA CIC ACA AGA AAA CIA CIA CIA ATT AAT CIA TIA CGA AGC ATC CIA CIA CAA AAC CAT 176 173 160

Gin Asp His Asn Asn Phe Gin Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC COO TAT GIT COO TGC AGT ACA TGT GAA GGT AAT CIT 200 186 182

Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro Thr Ile - C COA TOC TIA TOA CITC TOC OUT ATT GAG ACG GAA AGA COA COA AGC AGA COA COA ACA ATC - 31º Information for SEQ ID NO: 54 G198V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Cly Ser Ile Pro Pro Clu Asn His 5'- CIA COA GOC ACA AGA AAA COA COA ATT AAT COA TOA GGA AGO ATO COA COA GAA AAC CAT 160 173 176 Gin Asp His Asn Asn Phe Gin Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu CAA GEC CAC AAC AAC TIC CAA ACA CIC CCC TAT GIT CCC TGC AGT ACA TGT GAA GGT AAT CIT 182 186

Ala Cys Leu Ser Leu Cys His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro - C GCA TEC TIA TOA CITC TEC ONT ATT GRG ACE GRA AGA GCA COA ACE AGA GCA GCA - 3'

Information for SEQ ID NO: 55 G196V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single 10

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His 5'- CIA CCA CCC ACA AGA AAA CCA CCA ATT AAT CCA TCA CGA AGC ATC CCA CCA GAA AAC CAT 160 173 176 Gln Asp His Asn Asn The Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Glu Gly Asn Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC COO TAT GIT COO TGC AGT ACA TGT GAA GGT AAT CIT 182 **186** 196

Ala Cy : Leu Ser Leu Cys His Ille Glu Thr Glu Arg Ala Pro Ser Arg - C GCA TEC TEA TOA CITC 'L-C CAT AIT GAG ACE GAA AGA GCA COA AGC AGA - 3' Information for SEQ ID NO: 56 G194V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ille Asn Pro Ser Gly Ser Ille Pro Pro Glu Asn His 5'- CIA CCA CCI ACA AGA AAA CCA CCA ATT AAT CCA TCA CGA AGC ATC CCA CCA CAA AAC CAT 173 176

Gln Asp His Asn Asn Pre Gln The Leu Pro Tyr Val Pro Cys Ser The Cys Glu Gly Asn Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC CIC TAT GIT CIC TGC AGT ACA TGT GAA GGT AAT CIT 182 186

Ala Cys Leu Ser Leu Cys His Tle Clu Thr Clu Arg Ala Pro - C CCA TCC TTA TCA CTC TCC CAT ATT CAG ACG CAA ACA CCA CCA - 3'

Information for SEQ ID NO: 57 G192V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala The Arg Lys Pro Pro Ile Ash Pro See Gly See Ile Pro Pro Glu Ash His 5'- CIA CIA CIA CIA AGA AGA AGA CIA CIA CIA ATT AGT CIA TIA CIGA AGC ATC CIA CIA CIA CIA AGC 173 176 -

GIN ASP His ASN ASN FRE GIN The Leu Pro Tyr Val Pro Cys See The Cys Giu Gly ASN Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC CIC TAT GIT CIC TIC AGT ACA TGT GAA GGT AAT CIT 182 186 197

Ala Cys Leu Ser Leu Cys His The Glu Thr Glu Arg - C GCA TGC TTA TGA CTC TGC GAT ATT GGG AGG GAA Xxx - 3' Information for SEQ ID NO: 58 G6V

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single

TOPOLOGY: linear

## 5 MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His
5'- CTA CTA CTA CTC ACA ACA AAA CTA CCA ATT AAT CTA TCA CCA ACC ATC CTA CCA CAA AAC CAT
160

173

176

CIn Asp His Asn Asn Pre CIn Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys Clu Cly Asn Leu
CAA GC CAC AAC AAC TIC CAA ACA CTC CTC TAT GIT CTC TCC AGT ACA TGT GAA CGT AAT CTT
182

186

190

Ala Cys Leu Ser Leu Cys His Ile Glu Thr - C

Information for SEQ ID NO: 59 G7V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

10 TOPOLOGY: linear

MOLECULE TYPE: protein

158

N - Ash His Gin Asp His Ash Ash Phe Gin The Leu Pro Tyr Val. Pro Cys See The Cys
5'- Arc Cat Caa Gac Cac arc arc tic Caa ara cic cic tat Gi. cic tic agi aca tig

182

186

190

Glu Gly Ash Leu Ala Cys Leu See Leu Cys His Ile Glu The - C
Gaa Gif aat cit cia tat tic tia toa cic tic cat att gag acg - 3'

Information for SEQ ID NO: 60 G200V $\delta$ C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 61 amino acids, 183 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr. Arg Lys Pro Pro Ile Asn Pro Ser Cly Ser Ile Pro Pro Clu Asn His 5'- CIA CIA CIC ACA AGA AAA CIA CIA ATT AAT CIA TOA GGA AGC ATC CIA CIA GAA AAC CAT 160

Gin Asp His Asn Asn The Gin Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu CAA GCC CAC AAC AAC TIC CAA ACA CIC CIC TAT GIT CIC AGC AGT ACA TGT GAA CGT AAT CIT 182 186

Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro Thr Ile - C

Information for SEQ ID NO: 61 G198V $\delta$ C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 59 amino acids, 177 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Gly Ser Ile Pro Pro Glu Asn His 5'- CIA CIA GIC ACA AGA AAA CIA CIA CIA ATT AAT CIA TOA GGA AGC ATC CIA CIA GAA AAC CAT 160

Gin Asp His Asn Asn The Gin Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Giu Gly Asn Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC COC TAIT GIT COC AGC AGT ACA TGT GAA GGT AAT CIT 182 186

Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro Ser Arg Ala Pro - C

GCA TEC TTA TCA CIC AGE CAT ATT GAG AGE GAA AGA GCA CCA AGE AGA GCA CCA - 3'

Information for SEQ ID NO: 62 G196VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 57 amino acids, 171 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Cly Ser Ile Pro Pro Clu Asn His 5'- CIA CCA CCC ACA AGA AAA CCA CCA AIT AAT CCA TCA CGA AGC AIC CCA CCA CAA AAC CAT 160

GÎN ASP HIS ASN ASN THE GÎN THE LEU PRO TYE VAL PRO SEE SEE THE CYS GÎU GÎY ASN LEU CAA GAC CAC AAC AAC TIC CAA ACA CIC COC TAT GIT COC AGC AGT ACA TGT GAA CGT AAT CIT 182 186

Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro Ser Arg - C

GCA TEC TIA TCA CIC AGE CAT ATT GAG AGE GAA AGA GCA GCA AGE AGA - 3'

Information for SEQ ID NO: 63 G198VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 55 amino acids, 165 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

140

Gln Asp His Asn Asn Fre Gln Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Glu Gly Asn Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC CCC TAT GIT CCC AGC AGT ACA TGT GAA GGT AAT CIT 182 186

Ala Cys Leu Ser Leu Ser His Ile Glu Thr Glu Arg Ala Pro - C

GCA TGC TEA TCA CTC AGC CAT ATT GAG AGG GAA AGA GCA CCA - 3'

Information for SEQ ID NO: 64 G192VδC

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 52 amino acids, 156 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ile Asn Pro Ser Cily Ser Ile Pro Pro Cilu Asn His 5'- CIA CCA CCC ACA AGA AAA CCA CCA ATT AAT CCA TCA CGA AGC ATC CCA CCA GAA AAC CAT 160

GIN ASP His ASN ASN Five GIN The Leu Pro Tyr Val Pro See See The Cys Glu Gly ASN Leu CAA GAC CAC AAC AAC TIC CAA ACA CIC CIC TAT GIT CIC AGE AGT ACA TGT GAA GGT AAT CIT 182 186 192

Ala Cys Leu Ser Leu Ser His Ile Glu The Glu Arg - C

GCA TGC TTA TCA CTC AGC CAT ATT GAG AGG GAA AGA - 3'

Information for SEQ ID NO: 65 G6V&C

SEQUENCE LENGTH: 51 amino acids, 153 nucleotides

STRANDEDNESS: single

10 TOPOLOGY: linear

MOLECULE TYPE: protein

140

N - Leu Pro Ala Thr Arg Lys Pro Pro Ille Asn Pro Ser Gly Ser Ille Pro Pro Glu Asn His 5'- CIA CIA CIA ACA ACA AAA CIA CIA ATT AAT CIA TIA CIA ACC ATC CIA CIA CIA AAC CAT

160 173 176

Gir Asp His Ash Ash Phe Gin Thr Leu Pro Tyr Val Pro Ser Ser Thr Cys Giu Gly Ash Leu CAA GAC CAC AAC AAC TTC CAA ACA CTC CCC TAT GIT CCC AGC AGT ACA TGT GAA GGT AAT CTT 182 186 190

Ala Cys Leu Ser Leu Ser His Tie Glu Thr - C

GCA TGC TIA TCA CIT' AGC CAT AIT GAG AGG - 3'

Information for SEQ ID NO: 66 G7V&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 33 amino acids, 99 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

158

N - Ash His Gin Asp His Ash Ash Fhe Gin The Leu Pro Tyr Val Pro See See The Cys
5'- AFC CAT CAA GFC CAC AAC AAC TTC CAA ACA CTC CCC TAT GIT CCC AGC AGC AGT ACA TGT

182

186

190

Glu Gly Ash Leu Ala Cys Leu See Leu See His Lie Glu The - C

GAA GGT AAT CTT GCA TGC TTA TCA CTC AGC CAT ATT GFG AGG - 3'

Information for SEQ ID NO: 67 G4V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 51 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Cys Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His - C
5'- GTT CTC TGC AGT ACA TGT GAA GGT AAT CTT GCA TGC TTA TCA CTC TGC CAT - 3'

Information for SEQ ID NO: 68 G4V&C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 17 amino acids, 51 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187
N - Val Pro Ser Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His - C
5'- GTT CCC AGC AGT ACA TGT GAA GGT AAT CTT GCA TGC TTA TCA CTC AGC CAT - 3'

Information for SEQ ID NO: 69 G4'V

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 17 amino acids

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187 N - Val Pro Asp Ser Thr Asp Glu Gly Asn Leu Ala Orn Leu Ser Leu Orn His - C

Information for SEQ ID NO: 70 G4'VδC

SEQUENCE TYPE: amino acids

15 SEQUENCE LENGTH: 17 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

171 173 176 182 186 187 N - Val Pro Ser Ser Thr Asp Glu Gly Asn Leu Ala Orn Leu Ser Leu Ser His - C Information for SEQ ID NO: 71 G1V

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids, 42 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187

N - Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys His - C 5' - AGT ACA TGT GAA GGT AAT CTT GCA TGC TTA TCA CTC TGC CAT - 3'

Information for SEQ ID NO: 72 G1V $\delta$ C

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 14 amino acids,

STRANDEDNESS: single 10

TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182

N - Ser Thr Cys Glu Gly Asn Leu Ala Cys Leu Ser Leu Ser His - C

5'- AGT ACA TGT GAA GGT AAT CTT GCA TGC TTA TCA CTC AGC CAT - 3'

Information for SEQ ID NO: 73 G1'V&C

SEQUENCE TYPE: amino acids

SEQUENCE LENGTH: 14 amino acids

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

174 176 182 186 187 . N - Ser Thr Asp Glu Gly Asn Leu Ala Orn Leu Ser Leu Ser His - C

Information for SEQ ID NO: 74 BB

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 219 amino acids, 657 nucleotides

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

1

BB = N - Lys Tyr Gly Val Ser Asp Tyr Tyr Lys Asn Leu Ile Asn Asn Ala Lys 5'- AAA TAT GGA GIA AGT GAC TAT TAC AAG AAT CIA ATC AAC AAT CIC AAA

17

The Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val Val Glu Ser Ala Lys Lys ACT GIT GAA GCC GTA AAA GAC CTT CAA GCA CAA GTT GTT GAA TCA GCG AAG AAA

35

Ala Arg Ile Ser Glu Ala Thr Asp Gly Leu Ser Asp Phe Leu Lys Ser Glu Thr GUG GUT ATT TOA GAA GOA AOA GAT GUC TIA TOT GAT TIC TIG AAA TOA CAA AOA

53

Pro Ala Glu Asp The Val Lys See Lie Glu Leu Ala Glu Ala Lys Val Leu Ala CCT CCT GAA GAT ACT GIT AAA TCA ATT GAA TTA CCT GAA GCT AAA GTC TTA CCT 71

Ash Ang Giu Leu Asp Lys Tyr Gly Val Ser Asp Tyr His Lys Ash Leu Ile Ash AAC AGA GAA CIT GAC AAA TAT GGA GIA AGT GAC TAT CAC AAG AAC CIA ATC AAC

89

As Ala Lys Thr Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val Val Glu Ser AAT GCC AAA ACT GTT GAA GGT GTA AAA GAC CTT GAA GCA GAA GTT GTT GAA TCA

107

Ala Lys Lys Ala Arg Lle Ser Glu Ala Thr Asp Gly Leu Ser Asp Phe Leu Lys GCG AAG AAA GCG GGT ATT TCA GAA GCA ACA GAT GCC TTA TCT GAT TTC TTG AAA

125

Ser Gin The Pro Ala Giu Asp The Val Lys Ser Lie Giu Leu Ala Giu Ala Lys TCA CAA ACA CCT GCT GAA GCT ACT GTT AAA TCA ATT GAA TTA CCT GAA GCT AAA

143

Val Leu Ala Asn Arg Glu Leu Asp Lys Tyr Gly Val Ser Asp Tyr Tyr Lys Asn GTC TTA GCT AAC AGA GAA CTT GAC AAA TAT GGA GTA AGT GAC TAT TAC AAG AAC

161

Leu Ile Asn Asn Ala Lys Thr Val Glu Gly Val Lys Ala Leu Ile Asp Glu Ile CIA AIC AAC AAT GCC AAA ACT GTT GAA GGT GIA AAA GCA CIG ATA GAT GAA ATT

179

Leu Ala Ala Leu Pro Lys Thr Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr TIA GCT GCA TTA GCT AAG ACT GAC ACT TAC AAA TTA ATC CTT AAT GGT AAA ACA

197

LEU LYS GLY GLU THE THE GLU ALA VAL ASP ALA ALA THE ALA AND SEE THE TIG AAA GCC GAA ACA ACT ACT GAA GCT GTT GAT GCT GCT ACT GCA AGA TCT TIC

215

219

Asn the Pro 11e Leu - C
ANT TIC CCT ATC CTC - 3'

Information for SEQ ID NO: 75  $\delta$ BB = fragment of BB

SEQUENCE TYPE: amino acids and nucleotides

SEQUENCE LENGTH: 108 amino acids, 324 nucleotides

STRANDEDNESS: single

5 TOPOLOGY: linear

MOLECULE TYPE: protein

1

SEB = N -Lys Tyr Gly Val Ser Asp Tyr His Lys Asn Leu Ile Asn Asn Ala Lys 5'- AAA TAT GGA GIA AGT GGC TAT CAC AAG AAC CIA ATC AAC AAT GIC AAA 17

The Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val Val Glu See Ala Lys Lys Act GIT GAA GGT GIT GAA GGT GIT GAA GGT AAA GAC GTT GAA GGT GIT GAA TOA GGG AAA 35

Ala Arg Ile Ser Glu Ala Thr Asp Gly Leu Ser Asp Fhe Leu Lys Ser Gln Thr GCG CGT ATT TCA GAA GCA ACA GCT GCC TTA TCT GAT TTC TTG AAA TCA CAA ACA 53

Pro Ala Glu Asp The Val Lys See Ile Glu Leu Ala Glu Ala Lys Val Leu Ala CCT CCT GAA GAT ACT GTT AAA TCA ATT GAA TTA CCT GAA CCT AAA GTC TTA CCT 71

AST ATT GIU LEU ASP LYS TYR GIY VAL SER ASP TYR TYR LYS AST LEU TIE AST AAC AGA GAA CIT GAC AAA TAT GGA GTA AGT GAC TAT TAC AAG AAC CTA ATC AAC 89

AST Ala Lys The Val Cilu Cily Val Lys Ala Leu Ile Asp Cilu Ile Leu Ala Ala AAT CCC AAA ACT CIT CAA CCT CIA AAA CCA CIG ATA CAT CAA ATT TIA CCT CCA 107 108

Lau Pro - C

TIA CCT - 3'